Diagonal matrix

From Wikipedia, the free encyclopedia

In linear algebra, a **diagonal matrix** is a square matrix in which the entries outside the main diagonal (↘) are all zero. The diagonal entries themselves may or may not be zero. Thus, the matrix $D = (d_{i,j})$ with n columns and n rows is diagonal if:

$$d_{i,j} = 0 \text{ if } i \neq j \qquad \forall i,j \in \{1,2,\ldots,n\}.$$

For example, the following matrix is diagonal:

$$\begin{bmatrix} 1 & 0 & 0 \\ 0 & 4 & 0 \\ 0 & 0 & -3 \end{bmatrix}.$$

The term *diagonal matrix* may sometimes refer to a **rectangular diagonal matrix**, which is an m-by-n matrix with only the entries of the form $d_{i,i}$ possibly non-zero; for example,

$$\begin{bmatrix} 1 & 0 & 0 \\ 0 & 4 & 0 \\ 0 & 0 & -3 \\ 0 & 0 & 0 \end{bmatrix}, \text{ or } \begin{bmatrix} 1 & 0 & 0 & 0 & 0 \\ 0 & 4 & 0 & 0 & 0 \\ 0 & 0 & -3 & 0 & 0 \end{bmatrix}.$$

However, in the remainder of this article we will consider only square matrices.

Any diagonal matrix is also a symmetric matrix. Also, if the entries come from the field **R** or **C**, then it is a normal matrix as well.

Equivalently, we can define a diagonal matrix as a matrix that is both upper- and lower-triangular.

The identity matrix I_n and any square zero matrix are diagonal. A one-dimensional matrix is always diagonal.

Contents

Scalar matrix

A diagonal matrix with all its main diagonal entries equal is a **scalar matrix**, that is, a scalar multiple λI of the identity matrix I. Its effect on a vector is scalar multiplication by λ. For example, a 3×3 scalar matrix has the form:

$$\begin{bmatrix} \lambda & 0 & 0 \\ 0 & \lambda & 0 \\ 0 & 0 & \lambda \end{bmatrix}.$$

The scalar matrices are the center of the algebra of matrices: that is, they are precisely the matrices that commute with all other square matrices of the same size.

For an abstract vector space V (rather than the concrete vector space K^n), or more generally a module M over a ring R, with the endomorphism algebra End(M) (algebra of linear operators on M) replacing the algebra of matrices, the analog of scalar matrices are **scalar transformations**. Formally, scalar multiplication is a linear map, inducing a map $R \to \operatorname{End}(M)$, (send a scalar λ to the corresponding scalar transformation, multiplication by λ) exhibiting End(M) as a R-algebra. For vector spaces, or more generally free modules $M \cong R^n$, for which the endomorphism algebra is isomorphic to a matrix algebra, the scalar transforms are exactly the center of the endomorphism algebra, and similarly invertible transforms are the center of the general linear group GL(V), where they are denoted by Z(V), follow the usual notation for the center.

Matrix operations

The operations of matrix addition and matrix multiplication are especially simple for diagonal matrices. Write diag(a_1,...,a_n) for a diagonal matrix whose diagonal entries starting in the upper left corner are a_1,...,a_n. Then, for addition, we have

$$\operatorname{diag}(a_1,\ldots,a_n) + \operatorname{diag}(b_1,\ldots,b_n) = \operatorname{diag}(a_1+b_1,\ldots,a_n+b_n)$$

and for matrix multiplication,

$$\operatorname{diag}(a_1,\ldots,a_n) \cdot \operatorname{diag}(b_1,\ldots,b_n) = \operatorname{diag}(a_1 b_1,\ldots,a_n b_n).$$

The diagonal matrix diag(a_1,...,a_n) is invertible if and only if the entries a_1,...,a_n are all non-zero. In this case, we have

$$\operatorname{diag}(a_1,\ldots,a_n)^{-1} = \operatorname{diag}(a_1^{-1},\ldots,a_n^{-1}).$$

In particular, the diagonal matrices form a subring of the ring of all n-by-n matrices.

Multiplying an n-by-n matrix A from the *left* with diag(a_1,...,a_n) amounts to multiplying the i-th *row* of A by a_i for all i; multiplying the matrix A from the *right* with diag(a_1,...,a_n) amounts to multiplying the i-th *column* of A by a_i for all i.

Other properties

The eigenvalues of diag(a_1, ..., a_n) are a_1, ..., a_n with associated eigenvectors of e_1, ..., e_n, where the vector e_i is all zeros except a one in the ith row. The determinant of diag(a_1, ..., a_n) is the product $a_1 \ldots a_n$.

The adjugate of a diagonal matrix is again diagonal.

A square matrix is diagonal if and only if it is triangular and normal.

Uses

Diagonal matrices occur in many areas of linear algebra. Because of the simple description of the matrix operation and eigenvalues/eigenvectors given above, it is always desirable to represent a given matrix or linear map by a diagonal matrix.

In fact, a given n-by-n matrix A is similar to a diagonal matrix (meaning that there is a matrix X such that $X^{-1}AX$ is diagonal) if and only if it has n linearly independent eigenvectors. Such matrices are said to be diagonalizable.

Over the field of real or complex numbers, more is true. The spectral theorem says that every normal matrix is unitarily similar to a diagonal matrix (if $AA^* = A^*A$ then there exists a unitary matrix U such that UAU^* is diagonal). Furthermore, the singular value decomposition implies that for any matrix A, there exist unitary matrices U and V such that UAV^* is diagonal with positive entries.

Operator theory

In operator theory, particularly the study of PDEs, operators are particularly easy to understand, and PDEs easy to solve, if the operator is diagonal with respect to the basis one is working with – this corresponds to a separable partial differential equation. Thus, a key technique to understand operators is to have a change of coordinates – in the language of operators, an integral transform – which changes the basis to an eigenbasis of eigenfunctions: which makes the equation separable. An important example of this is the Fourier transform, which diagonalizes the heat equation.

See also

- Anti-diagonal matrix
- Banded matrix

- Bidiagonal matrix
- Diagonally dominant matrix
- Diagonalizable matrix
- Tridiagonal matrix
- Toeplitz matrix
- Toral Lie algebra
- Circulant matrix

References

- Roger A. Horn and Charles R. Johnson, *Matrix Analysis*, Cambridge University Press, 1985. ISBN 0-521-30586-1 (hardback), ISBN 0-521-38632-2 (paperback).

Categories: Matrices | Matrix normal forms

Rotation matrix

From Wikipedia, the free encyclopedia

In linear algebra, a **rotation matrix** is any matrix that acts as a rotation in Euclidean space. For example, the matrix

$$\begin{bmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{bmatrix}$$

rotates vectors in the xy-plane counterclockwise by an angle of θ. In three dimensions, rotation matrices are among the simplest algebraic descriptions of rotations, and are used extensively for computations in geometry, physics, and computer graphics.

Though most applications involve rotations in 2 or 3 dimensions, rotation matrices can be defined for n-dimensional space. Algebraically, a rotation matrix is an orthogonal matrix whose determinant is equal to 1:

$$R^T = R^{-1} \quad \text{and} \quad \det R = 1.$$

Rotation matrices are always square, and are usually assumed to have real entries, though the definition makes sense for other scalar fields. The set of all $n \times n$ rotation matrices forms a group, known as the rotation group (or special orthogonal group).

Contents

Rotations in two and three dimensions

In all of this section, the matrices are assumed to act on column vectors, for instance in cartesian coordinates systems: [x,y,z] in 3D or [x,y] in 2D (each of them transposed to a column vector).

Dimension two

In two dimensions, every rotation matrix has the following form:

$$R(\theta) = \begin{bmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{bmatrix} \quad (\text{rotation by } \theta).$$

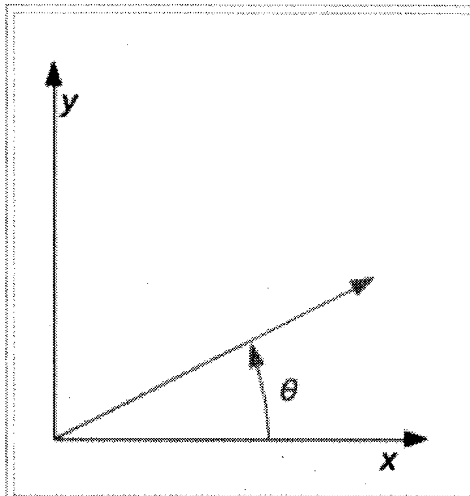
This matrix rotates the plane around the origin by an angle of θ. The x axis is rotated towards the y axis.

The new coordinates (x',y') for a point (x,y) will then be given by:

$$x' = x\cos\theta - y\sin\theta$$
$$y' = x\sin\theta + y\cos\theta$$

In an oriented plane

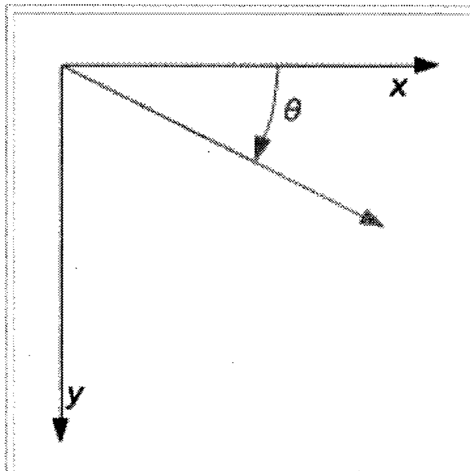
If we use the standard right-handed coordinate system, where x axis goes to the right and where y axis goes up, the rotation is counterclockwise. If one uses the opposite convention, for example x directed to the right and y directed to the bottom, the rotation will be clockwise. To get convinced that it is still the same rotation, one can think at the plane as a sheet of paper being watched alternatively from above and from beneath, by transparency.



The effect of the rotation matrix in a plane oriented in the standard way

Such non-standard orientations are almost never used in mathematics and physics, but they are very common in computer graphics[1], because they match the direction of writing for Western scripts: from the left to the right and from the top to the bottom. That's the reason why, in much computer software, rotations go clockwise.

Assuming the standard orientation, for a clockwise rotation, simply replace θ by $-\theta$:



The effect of the same rotation matrix in a plane that is oriented otherwise

$$R(\theta) = \begin{bmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{bmatrix} \quad \text{(counterclockwise rotation by } \theta\text{)}.$$

$$R(-\theta) = \begin{bmatrix} \cos\theta & \sin\theta \\ -\sin\theta & \cos\theta \end{bmatrix} \quad \text{(clockwise rotation by } \theta\text{)}.$$

Common rotations

Particularly useful are the matrices for 90° and 180° rotations:

$$R(90^\circ) = \begin{bmatrix} 0 & -1 \\ 1 & 0 \end{bmatrix} \quad \text{(90° counterclockwise rotation)}.$$

$$R(180^\circ) = \begin{bmatrix} -1 & 0 \\ 0 & -1 \end{bmatrix} \quad \text{(180° rotation)}.$$

$$R(270^\circ) = \begin{bmatrix} 0 & 1 \\ -1 & 0 \end{bmatrix} \quad \text{(90° clockwise rotation)}.$$

Dimension three

See also: Rotation representation

Basic rotations

There are three basic rotation matrices in three dimensions:

$$R_x(\theta) = \begin{bmatrix} 1 & 0 & 0 \\ 0 & \cos\theta & -\sin\theta \\ 0 & \sin\theta & \cos\theta \end{bmatrix}$$

$$R_y(\theta) = \begin{bmatrix} \cos\theta & 0 & \sin\theta \\ 0 & 1 & 0 \\ -\sin\theta & 0 & \cos\theta \end{bmatrix}$$

$$R_z(\theta) = \begin{bmatrix} \cos\theta & -\sin\theta & 0 \\ \sin\theta & \cos\theta & 0 \\ 0 & 0 & 1 \end{bmatrix}.$$

These matrices represent counterclockwise rotations of an object relative to fixed coordinate axes, by an angle of θ. The direction of the rotation is as follows: R_x rotates the y-axis towards the z-axis, R_y rotates the z-axis towards the x-axis, and R_z rotates the x-axis towards the y-axis.

The resulting coordinates (x',y',z') for a point (x,y,z), for each of these rotations are:

	x'	y'	z'
$\mathbf{R_x(\theta)}$	x	$y\cos\theta - z\sin\theta$	$z\cos\theta + y\sin\theta$
$\mathbf{R_y(\theta)}$	$x\cos\theta + z\sin\theta$	y	$-x\sin\theta + z\cos\theta$

$\mathbf{R_z(\theta)}$	$x\cos\theta - y\sin\theta$	$y\cos\theta + x\sin\theta$	z

In an oriented space

If the 3D space is oriented in the usual way, i.e. x going to the right, y going to the front and z going up, these three rotations are counterclockwise when the third (unchanged) axis goes towards the observer. This direction of the rotation can be determined by the right-hand rule.

General rotations

Other rotation matrices can be obtained from these three using matrix multiplication. For example, the product

$$R_x(\gamma)\,R_y(\beta)\,R_z(\alpha)$$

represents a rotation whose yaw, pitch, and roll are α, β, and γ, respectively. Similarly, the product

$$R_z(\gamma)\,R_x(\beta)\,R_z(\alpha)$$

represents a rotation whose Euler angles are α, β, and γ (using the **z**-**x**-**z** convention for Euler angles).

Finding the rotation matrix

Every rotation in three dimensions is defined by its **axis** — a direction that is left fixed by the rotation — and its **angle** — the amount by which the rotation turns.

Determining the axis

Given a rotation matrix R, a vector $\mathbf{u}$ parallel to the rotation axis must satisfy

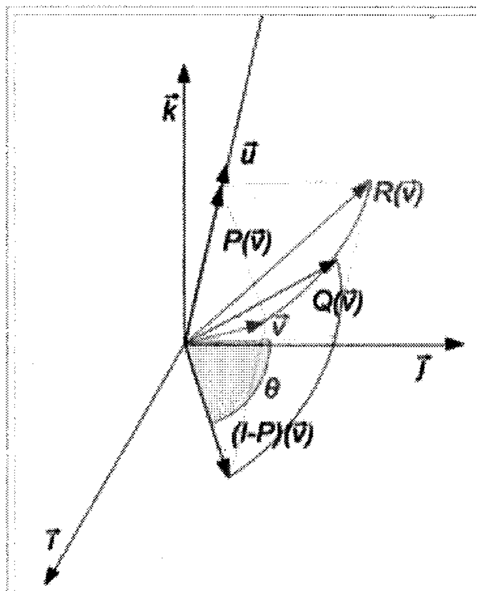
$$R\mathbf{u} = \mathbf{u}$$

since the rotation of $\mathbf{u}$ around the rotation axis must result in $\mathbf{u}$. The equation above may be solved for $\mathbf{u}$ which is unique up to a scalar factor.

Further, the equation may be rewritten

$$R\mathbf{u} = I\mathbf{u} \quad \Rightarrow \quad (R - I)\mathbf{u} = 0$$

which shows that $\mathbf{u}$ is the null space of $R - I$. Viewed another way, $\mathbf{u}$ is an eigenvector corresponding to the eigenvalue $\lambda = 1$ (every rotation matrix must have this eigenvalue).



A rotation **R** around axis **u** can be decomposed using 3 endomorphisms **P**, **(I - P)**, and **Q** (click to enlarge).

Determining the angle

To find the angle of a rotation, once the axis of the rotation is known, select a vector $\mathbf{v}$ perpendicular to the axis. Then the angle of the rotation is the

angle between $\mathbf{v}$ and $R\mathbf{v}$.

Rotation matrix given an axis and an angle

For some applications, it is helpful to be able to make a rotation with a given axis. Given a unit vector $\mathbf{u} = (u_x, u_y, u_z)$, where $u_x^2 + u_y^2 + u_z^2 = 1$, the matrix for a rotation by an angle of θ about an axis in the direction of $\mathbf{u}$ is[2]:

$$R = \begin{bmatrix} u_x^2 + (1-u_x^2)c & u_x u_y(1-c) - u_z s & u_x u_z(1-c) + u_y s \\ u_x u_y(1-c) + u_z s & u_y^2 + (1-u_y^2)c & u_y u_z(1-c) - u_x s \\ u_x u_z(1-c) - u_y s & u_y u_z(1-c) + u_x s & u_z^2 + (1-u_z^2)c \end{bmatrix},$$

where

$$c = \cos\theta, \quad s = \sin\theta.$$

This can be written more concisely as

$$R = \mathbf{u}\otimes\mathbf{u} + \cos\theta(I - \mathbf{u}\otimes\mathbf{u}) + \sin\theta[\mathbf{u}]_\times,$$

where $[\mathbf{u}]_\times$ is the skew symmetric form of $\mathbf{u}$, and $\otimes$ is the outer product.

If the 3D space is oriented in the usual way, this rotation will be counterclockwise for an observer placed so that the axis $\mathbf{u}$ goes in his or her direction (Right-hand rule).

Simpler form of the axis-angle formula

Rodrigues' rotation formula can be written as

$$R = P + (I - P)\cos\theta + Q\sin\theta$$

where

$$P = \begin{bmatrix} u_x^2 & u_x u_y & u_x u_z \\ u_x u_y & u_y^2 & u_y u_z \\ u_x u_z & u_y u_z & u_z^2 \end{bmatrix} = \mathbf{u}\mathbf{u}^T, \qquad I = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}, \qquad Q = \begin{bmatrix} 0 & -u_z & u_y \\ u_z & 0 & -u_x \\ -u_y & u_x & 0 \end{bmatrix}.$$

The matrix I is the 3×3 identity matrix. The matrix Q is the skew-symmetric representation of a cross product with $\mathbf{u}$. The matrix P is the projection onto the axis of rotation, and $I - P$ is the projection onto the plane orthogonal to the axis.

Properties of a rotation matrix

The above discussion can be generalised to any number of dimensions. For any rotation matrix $R_{axis,\theta} \in \mathbb{R}^n$ and

I, the identity in $I \in \mathbb{R}^n$

$$\begin{aligned} &R_{a,(\theta+r)} = R_{a,\theta} \cdot R_{a,r} \\ &R_{a,0} = \mathbb{I} \\ &R^T_{a,\theta} = R^{-1}_{a,\theta} \\ &\det\left(R_{a,\theta}\right) = 1 \end{aligned}$$

Examples

- The 2×2 rotation matrix

$$Q = \begin{bmatrix} 0 & -1 \\ 1 & 0 \end{bmatrix}$$

corresponds to a 90° planar rotation.
- The transpose of the 2×2 matrix

$$M = \begin{bmatrix} 0.936 & 0.352 \\ 0.352 & -0.936 \end{bmatrix}$$

is its inverse, but since its determinant is −1 this is not a rotation matrix; it is a reflection across the line $11y = 2x$.
- The 3×3 rotation matrix

$$Q = \begin{bmatrix} 1 & 0 & 0 \\ 0 & \frac{\sqrt{3}}{2} & \frac{1}{2} \\ 0 & -\frac{1}{2} & \frac{\sqrt{3}}{2} \end{bmatrix}$$

corresponds to a −30° rotation around the x axis in three-dimensional space.
- The 3×3 rotation matrix

$$Q = \begin{bmatrix} 0.36 & 0.48 & -0.8 \\ -0.8 & 0.60 & 0 \\ 0.48 & 0.64 & 0.60 \end{bmatrix}$$

corresponds to a rotation of approximately 74° around the axis $(-\frac{1}{3},\frac{2}{3},\frac{2}{3})$ in three-dimensional space.
- The 3×3 permutation matrix

$$\begin{bmatrix} 0 & 0 & 1 \end{bmatrix}$$

- The 3×3 matrix

$$M = \begin{bmatrix} 3 & -4 & 1 \\ 5 & 3 & -7 \\ -9 & 2 & 6 \end{bmatrix}$$

has determinant +1, but its transpose is not its inverse, so it is not a rotation matrix.
- The 4×3 matrix

$$M = \begin{bmatrix} 0.5 & -0.1 & 0.7 \\ 0.1 & 0.5 & -0.5 \\ -0.7 & 0.5 & 0.5 \\ -0.5 & -0.7 & -0.1 \end{bmatrix}$$

is not square, and so cannot be a rotation matrix; yet M^TM yields a 3×3 identity matrix (the columns are orthonormal).
- The 4×4 matrix

$$Q = \begin{bmatrix} -1 & 0 & 0 & 0 \\ 0 & -1 & 0 & 0 \\ 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & -1 \end{bmatrix}$$

describes an isoclinic rotation, a rotation through equal angles (180°) through two orthogonal planes.
- The 5×5 rotation matrix

$$Q = \begin{bmatrix} 0 & -1 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & -1 & 0 & 0 \\ 0 & 0 & 0 & -1 & 0 \end{bmatrix}$$

$$P = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \end{bmatrix}$$

is a rotation matrix, as is the matrix of any even permutation, and rotates through 120° about the axis $x = y = z$.

$$\begin{bmatrix} 0 & 0 & 0 & 0 & 1 \end{bmatrix}$$

rotates vectors in the plane of the first two coordinate axes 90°, rotates vectors in the plane of the next two axes 180°, and leaves the last coordinate axis unmoved.

Geometry

In Euclidean geometry, a rotation is an example of an isometry, a transformation that moves points without changing the distances between them. Rotations are distinguished from other isometries by two additional properties: they leave (at least) one point fixed, and they leave "handedness" unchanged. By contrast, a translation moves every point, a reflection exchanges left- and right-handed ordering, and a glide reflection does both.

A rotation that does not leave "handedness" unchanged is called an Improper Rotation or a Rotoinversion

If we take the fixed point as the origin of a Cartesian coordinate system, then every point can be given coordinates as a displacement from the origin. Thus we may work with the vector space of displacements instead of the points themselves. Now suppose $(p_1,\ldots,p_n)$ are the coordinates of the vector **p** from the origin, O, to point P. Choose an orthonormal basis for our coordinates; then the squared distance to P, by Pythagoras, is

$$d^2(O,P) = \|\mathbf{p}\|^2 = \sum_{r=1}^{n} p_r^2$$

which we can compute using the matrix multiplication

$$\|\mathbf{p}\|^2 = \begin{bmatrix} p_1 \cdots p_n \end{bmatrix} \begin{bmatrix} p_1 \\ \vdots \\ p_n \end{bmatrix} = \mathbf{p}^T\mathbf{p}.$$

A geometric rotation transforms lines to lines, and preserves ratios of distances between points. From these properties we can show that a rotation is a linear transformation of the vectors, and thus can be written in matrix form, $Q\mathbf{p}$. The fact that a rotation preserves, not just ratios, but distances themselves, we can state as

$$\mathbf{p}^T\mathbf{p} = (Q\mathbf{p})^T(Q\mathbf{p}),$$

or

$$\begin{aligned} \mathbf{p}^T I \mathbf{p} &= (\mathbf{p}^T Q^T)(Q\mathbf{p}) \\ &= \mathbf{p}^T (Q^T Q)\mathbf{p}. \end{aligned}$$

Because this equation holds for all vectors, **p**, we conclude that every rotation matrix, Q, satisfies the *orthogonality* condition,

$$Q^T Q = I.$$

Rotations preserve handedness because they cannot change the ordering of the axes, which implies the *special matrix* condition,

$$\det Q = +1.$$

Equally important, we can show that any matrix satisfying these two conditions acts as a rotation.

Multiplication

The inverse of a rotation matrix is its transpose, which is also a rotation matrix:

$$\begin{aligned}(Q^T)^T(Q^T) &= QQ^T = I \\ \det Q^T &= \det Q = +1.\end{aligned}$$

The product of two rotation matrices is a rotation matrix:

$$\begin{aligned}(Q_1Q_2)^T(Q_1Q_2) &= Q_2^T(Q_1^TQ_1)Q_2 = I \\ \det(Q_1Q_2) &= (\det Q_1)(\det Q_2) = +1.\end{aligned}$$

For n greater than 2, multiplication of $n \times n$ rotation matrices is not commutative.

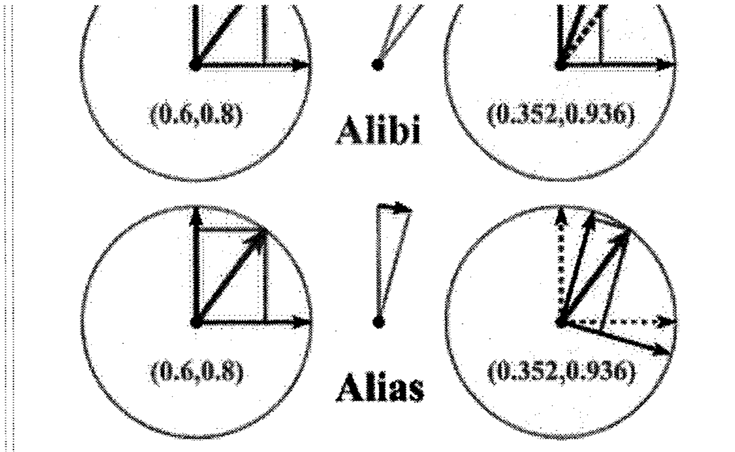
$$Q_1 = \begin{bmatrix} 0 & -1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 1 \end{bmatrix} \qquad Q_2 = \begin{bmatrix} 0 & 0 & 1 \\ 0 & 1 & 0 \\ -1 & 0 & 0 \end{bmatrix}$$

$$Q_1Q_2 = \begin{bmatrix} 0 & -1 & 0 \\ 0 & 0 & 1 \\ -1 & 0 & 0 \end{bmatrix} \quad Q_2Q_1 = \begin{bmatrix} 0 & 0 & 1 \\ 1 & 0 & 0 \\ 0 & 1 & 0 \end{bmatrix}.$$

Noting that any identity matrix is a rotation matrix, and that matrix multiplication is associative, we may summarize all these properties by saying that the $n \times n$ rotation matrices form a group, which for $n > 2$ is non-abelian. Called a special orthogonal group, and denoted by SO(n), SO(n,**R**), SO$_n$, or SO$_n$(**R**), the group of $n \times n$ rotation matrices is isomorphic to the group of rotations in an n-dimensional space. This means that multiplication of rotation matrices corresponds to composition of rotations, applied in left-to-right order of their corresponding matrices.

Ambiguities

The interpretation of a rotation matrix can be subject to many ambiguities.



Alias and alibi rotations

Alias or alibi transformation
: The change in a vector's coordinates can indicate a turn of the coordinate system (alias) or a turn of the vector (alibi).

Right- or left-handed coordinates
: The matrix can be with respect to a right-handed or left-handed coordinate system.

World or body axes
: The coordinate axes can be fixed or rotate with a body.

Vectors or forms
: The vector space has a dual space of linear forms, and the matrix can act on either vectors or forms.

In most cases the effect of the ambiguity is to transpose or invert the matrix.

Decompositions

Independent planes

Consider the 3×3 rotation matrix

$$Q = \begin{bmatrix} 0.36 & 0.48 & -0.8 \\ -0.8 & 0.60 & 0 \\ 0.48 & 0.64 & 0.60 \end{bmatrix}.$$

If Q acts in a certain direction, $\mathbf{v}$, purely as a scaling by a factor λ, then we have

$$Q\mathbf{v} = \lambda\mathbf{v},$$

so that

$$\mathbf{0} = (\lambda I - Q)\mathbf{v}.$$

Thus λ is a root of the characteristic polynomial for Q,

$$\begin{aligned} 0 &= \det(\lambda I - Q) \\ &= \lambda^3 - \tfrac{39}{25}\lambda^2 + \tfrac{39}{25}\lambda - 1 \\ &= (\lambda - 1)(\lambda^2 - \tfrac{14}{25}\lambda + 1). \end{aligned}$$

Two features are noteworthy. First, one of the roots (or eigenvalues) is 1, which tells us that some direction is unaffected by the matrix. For rotations in three dimensions, this is the *axis* of the rotation (a concept that has no meaning in any other dimension). Second, the other two roots are a pair of complex conjugates, whose product is 1 (the constant term of the quadratic), and whose sum is 2 cos θ (the negated linear term). This factorization is of interest for 3×3 rotation matrices because the same thing occurs for all of them. (As special cases, for a null rotation the "complex conjugates" are both 1, and for a 180° rotation they are both −1.) Furthermore, a similar factorization

holds for any $n \times n$ rotation matrix. If the dimension, n, is odd, there will be a "dangling" eigenvalue of 1; and for any dimension the rest of the polynomial factors into quadratic terms like the one here (with the two special cases noted). We are guaranteed that the characteristic polynomial will have degree n and thus n eigenvalues. And since a rotation matrix commutes with its transpose, it is a normal matrix, so can be diagonalized. We conclude that every rotation matrix, when expressed in a suitable coordinate system, partitions into independent rotations of two-dimensional subspaces, at most $^n/_2$ of them.

The sum of the entries on the main diagonal of a matrix is called the trace; it does not change if we reorient the coordinate system, and always equals the sum of the eigenvalues. This has the convenient implication for 2×2 and 3×3 rotation matrices that the trace reveals the angle of rotation, θ, in the two-dimensional (sub-)space. For a 2×2 matrix the trace is 2 cos(θ), and for a 3×3 matrix it is 1+2 cos(θ). In the three-dimensional case, the subspace consists of all vectors perpendicular to the rotation axis (the invariant direction, with eigenvalue 1). Thus we can extract from any 3×3 rotation matrix a rotation axis and an angle, and these completely determine the rotation.

Sequential angles

The constraints on a 2×2 rotation matrix imply that it must have the form

$$Q = \begin{bmatrix} a & -b \\ b & a \end{bmatrix}$$

with $a^2+b^2 = 1$. Therefore we may set $a = \cos\theta$ and $b = \sin\theta$, for some angle θ. To solve for θ it is not enough to look at a alone or b alone; we must consider both together to place the angle in the correct quadrant, using a two-argument arctangent function.

Now consider the first column of a 3×3 rotation matrix,

$$\begin{bmatrix} a \\ b \\ c \end{bmatrix}.$$

Although a^2+b^2 will probably not equal 1, but some value $r^2 < 1$, we can use a slight variation of the previous computation to find a so-called Givens rotation that transforms the column to

$$\begin{bmatrix} r \\ 0 \\ c \end{bmatrix},$$

zeroing b. This acts on the subspace spanned by the x and y axes. We can then repeat the process for the xz subspace to zero c. Acting on the full matrix, these two rotations produce the schematic form

$$Q_{xz}Q_{xy}Q = \begin{bmatrix} 1 & 0 & 0 \\ 0 & * & * \\ 0 & * & * \end{bmatrix}.$$

Shifting attention to the second column, a Givens rotation of the yz subspace can now zero the z value. This brings the full matrix to the form

$$Q_{yz}Q_{xz}Q_{xy}Q = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix},$$

which is an identity matrix. Thus we have decomposed Q as

An $n \times n$ rotation matrix will have $(n-1)+(n-2)+\cdots+2+1$, or

entries below the diagonal to zero. We can zero them by extending the same idea of stepping through the columns with a series of rotations in a fixed sequence of planes. We conclude that the set of $n \times n$ rotation matrices, each of which has n^2 entries, can be parameterized by $n(n-1)/2$ angles.

xzx_w	xzy_w	xyx_w	xyz_w
yxy_w	yxz_w	yzy_w	yzx_w
zyz_w	zyx_w	zxz_w	zxy_w
xzx_b	yzx_b	xyx_b	zyx_b
yxy_b	zxy_b	yzy_b	xzy_b
zyz_b	xyz_b	zxz_b	yxz_b

In three dimensions this restates in matrix form an observation made by Euler, so mathematicians call the ordered sequence of three angles Euler angles. However, the situation is somewhat more complicated than we have so far indicated. Despite the small dimension, we actually have considerable freedom in the sequence of axis pairs we use; and we also have some freedom in the choice of angles. Thus we find many different conventions employed when three-dimensional rotations are parameterized for physics, or medicine, or chemistry, or other disciplines. When we include the option of world axes or body axes, 24 different sequences are possible. And while some disciplines call any sequence Euler angles, others give different names (Euler, Cardano, Tait-Bryan, roll-pitch-yaw) to different sequences.

One reason for the large number of options is that, as noted previously, rotations in three dimensions (and higher) do not commute. If we reverse a given sequence of rotations, we get a different outcome. This also implies that we cannot compose two rotations by adding their corresponding angles. Thus *Euler angles are not vectors*, despite a similarity in appearance as a triple of numbers.

Nested dimensions

A 3×3 rotation matrix like

suggests a 2×2 rotation matrix,

is embedded in the upper left corner:

This is no illusion; not just one, but many, copies of n-dimensional rotations are found within $(n+1)$-dimensional rotations, as subgroups. Each embedding leaves one direction fixed, which in the case of 3×3 matrices is the rotation axis. For example, we have

fixing the x axis, the y axis, and the z axis, respectively. The rotation axis need not be a coordinate axis; if $\mathbf{u} = (x,y,z)$ is a unit vector in the desired direction, then

where $c_\theta = \cos\theta$, $s_\theta = \sin\theta$, is a rotation by angle θ leaving axis $\mathbf{u}$ fixed.

A direction in (n+1)-dimensional space will be a unit magnitude vector, which we may consider a point on a generalized sphere, S^n. Thus it is natural to describe the rotation group SO(n+1) as combining SO(n) and S^n. A suitable formalism is the fiber bundle,

where for every direction in the "base space", S^n, the "fiber" over it in the "total space", SO(n+1), is a copy of the "fiber space", SO(n), namely the rotations that keep that direction fixed.

Thus we can build an $n\times n$ rotation matrix by starting with a 2×2 matrix, aiming its fixed axis on S^2 (the ordinary sphere in three-dimensional space), aiming the resulting rotation on S^3, and so on up through S^{n-1}. A point on S^n can be selected using n numbers, so we again have $n(n-1)/2$ numbers to describe any $n\times n$ rotation matrix.

In fact, we can view the sequential angle decomposition, discussed previously, as reversing this process. The composition of $n-1$ Givens rotations brings the first column (and row) to (1,0,…,0), so that the remainder of the matrix is a rotation matrix of dimension one less, embedded so as to leave (1,0,…,0) fixed.

Skew parameters via Cayley's formula

When an $n\times n$ rotation matrix, Q, does not include -1 as an eigenvalue, so that none of the planar rotations of which it is composed are 180° rotations, then $Q+I$ is an invertible matrix. Most rotation matrices fit this description, and for them we can show that $(Q-I)(Q+I)^{-1}$ is a skew-symmetric matrix, A. Thus $A^{\mathrm{T}} = -A$; and since the diagonal is necessarily zero, and since the upper triangle determines the lower one, A contains $n(n-1)/2$ independent numbers. Conveniently, $I-A$ is invertible whenever A is skew-symmetric; thus we can recover the original matrix using the *Cayley transform*,

which maps any skew-symmetric matrix A to a rotation matrix. In fact, aside from the noted exceptions, we can produce any rotation matrix in this way. Although in practical applications we can hardly afford to ignore 180° rotations, the Cayley transform is still a potentially useful tool, giving a parameterization of most rotation matrices without trigonometric functions.

In three dimensions, for example, we have (Cayley 1846)

If we condense the skew entries into a vector, (x,y,z), then we produce a 90° rotation around the x axis for (1,0,0), around the y axis for (0,1,0), and around the z axis for (0,0,1). The 180° rotations are just out of reach; for, in the limit as x goes to infinity, $(x,0,0)$ does approach a 180° rotation around the x axis, and similarly for other directions.

Lie theory

Lie group

We have established that $n \times n$ rotation matrices form a group, the special orthogonal group, SO(n). This algebraic structure is coupled with a topological structure, in that the operations of multiplication and taking the inverse (which here is merely transposition) are continuous functions of the matrix entries. Thus SO(n) is a classic example of a topological group. (In purely topological terms, it is a compact manifold.) Furthermore, the operations are not only continuous, but smooth, so SO(n) is a differentiable manifold and a Lie group (Baker (2003); Fulton & Harris (1991)).

Most properties of rotation matrices depend very little on the dimension, n; yet in Lie group theory we see systematic differences between even dimensions and odd dimensions. As well, there are some irregularities below $n = 5$; for example, SO(4) is, anomalously, not a simple Lie group, but instead isomorphic to the product of S^3 and SO(3).

Lie algebra

Associated with every Lie group is a Lie algebra, a linear space equipped with a bilinear alternating product called a bracket. The algebra for SO(n) is denoted by

and consists of all skew-symmetric $n \times n$ matrices (as implied by differentiating the orthogonality condition, $I = Q^T Q$). The bracket, $[A_1, A_2]$, of two skew-symmetric matrices is defined to be $A_1 A_2 - A_2 A_1$, which is again a skew-symmetric matrix. This Lie algebra bracket captures the essence of the Lie group product via infinitesimals.

For 2×2 rotation matrices, the Lie algebra is a one-dimensional vector space, multiples of

Here the bracket always vanishes, which tells us that, in two dimensions, rotations commute. Not so in any higher dimension. For 3×3 rotation matrices, we have a three-dimensional vector space with the convenient basis

The Lie brackets of these generators are as follows

We can conveniently identify any matrix in this Lie algebra with a vector in $\mathbf{R}^3$,

Under this identification, the **so**(3) bracket has a memorable description; it is the vector cross product,

The matrix identified with a vector **v** is also memorable, because

Notice this implies that **v** is in the null space of the skew-symmetric matrix with which it is identified, because $\mathbf{v} \times \mathbf{v}$ is always the zero vector.

Exponential map

Connecting the Lie algebra to the Lie group is the *exponential map*, which we define using the familiar power series for e^x (Wedderburn 1934, §8.02),

For any skew-symmetric A, exp(A) is always a rotation matrix.

An important practical example is the 3×3 case, where we have seen we can identify every skew-symmetric matrix with a vector $\boldsymbol{\omega} = \mathbf{u}\theta$, where $\mathbf{u} = (x,y,z)$ is a unit magnitude vector. Recall that $\mathbf{u}$ is in the null space of the matrix associated with $\boldsymbol{\omega}$, so that if we use a basis with $\mathbf{u}$ as the z axis the final column and row will be zero. Thus we know in advance that the exponential matrix must leave $\mathbf{u}$ fixed. It is mathematically impossible to supply a straightforward formula for such a basis as a function of $\mathbf{u}$ (its existence would violate the hairy ball theorem), but direct exponentiation is possible, and yields

where $c = \cos {}^{\theta}\!/_{2}$, $s = \sin {}^{\theta}\!/_{2}$. We recognize this as our matrix for a rotation around axis $\mathbf{u}$ by angle θ. We also note that this mapping of skew-symmetric matrices is quite different from the Cayley transform discussed earlier.

In any dimension, if we choose some nonzero A and consider all its scalar multiples, exponentiation yields rotation matrices along a *geodesic* of the group manifold, forming a one-parameter subgroup of the Lie group. More broadly, the exponential map provides a homeomorphism between a neighborhood of the origin in the Lie algebra and a neighborhood of the identity in the Lie group. In fact, we can produce any rotation matrix as the exponential of some skew-symmetric matrix, so for these groups the exponential map is a *surjection*.

Baker–Campbell–Hausdorff formula

Suppose we are given A and B in the Lie algebra. Their exponentials, exp(A) and exp(B), are rotation matrices, which we can multiply. Since the exponential map is a surjection, we know that for some C in the Lie algebra, exp(A)exp(B) = exp(C), and we write

When exp(A) and exp(B) commute (which always happens for 2×2 matrices, but not higher), then $C = A+B$, mimicking the behavior of complex exponentiation. The general case is given by the BCH formula, a series expanded in terms of the bracket (Hall 2004, Ch. 3; Varadarajan 1984, §2.15). For matrices, the bracket is the same operation as the commutator, which detects lack of commutativity in multiplication. The general formula begins as follows.

Representation of a rotation matrix as a sequential angle decomposition, as in Euler angles, may tempt us to treat rotations as a vector space, but the higher order terms in the BCH formula reveal that to be a mistake.

We again take special interest in the 3×3 case, where $[A,B]$ equals the cross product, $A\times B$. If A and B are linearly independent, then A, B, and $A\times B$ can be used as a basis; if not, then A and B commute. And conveniently, in this dimension the summation in the BCH formula has a closed form (Engø 2001) as $\alpha A+\beta B+\gamma(A\times B)$.

Spin group

We again take special interest in the 3×3 case, where $[A,B]$ equals the cross product, $A \times B$. If A and B are linearly independent, then A, B, and $A \times B$ can be used as a basis; if not, then A and B commute. And conveniently, in this dimension the summation in the BCH formula has a closed form (Engø 2001) as $\alpha A+\beta B+\gamma(A \times B)$.

Spin group

The Lie group of $n \times n$ rotation matrices, SO(n), is a compact and path-connected manifold, and thus locally compact and connected. However, it is not simply connected, so Lie theory tells us it is a kind of "shadow" (a homomorphic image) of a universal covering group. Often the covering group, which in this case is the spin group denoted by Spin(n), is simpler and more natural to work with (Baker 2003, Ch. 5; Fulton & Harris 1991, pp. 299–315).

In the case of planar rotations, SO(2) is topologically a circle, S^1. Its universal covering group, Spin(2), is isomorphic to the real line, $\mathbf{R}$, under addition. In other words, whenever we use angles of arbitrary magnitude, which we often do, we are essentially taking advantage of the convenience of the "mother space". Every 2×2 rotation matrix is produced by a countable infinity of angles, separated by integer multiples of 2π. Correspondingly, the fundamental group of SO(2) is isomorphic to the integers, $\mathbf{Z}$.

In the case of spatial rotations, SO(3) is topologically equivalent to three-dimensional real projective space, $\mathbf{RP}^3$. Its universal covering group, Spin(3), is isomorphic to the 3-sphere, S^3. Every 3×3 rotation matrix is produced by two opposite points on the sphere. Correspondingly, the fundamental group of SO(2) is isomorphic to the two-element group, $\mathbf{Z}_2$. We can also describe Spin(3) as isomorphic to quaternions of unit norm under multiplication, or to certain 4×4 real matrices, or to 2×2 complex special unitary matrices.

Concretely, a unit quaternion, q, with

produces the rotation matrix

This is our third version of this matrix, here as a rotation around non-unit axis vector (x,y,z) by angle 2θ, where $\cos\theta = w$ and $|\sin\theta| = \|(x,y,z)\|$. (The proper sign for $\sin\theta$ is implied once the signs of the axis components are decided.)

Many features of this case are the same for higher dimensions. The coverings are all two-to-one, with SO(n), $n > 2$, having fundamental group $\mathbf{Z}_2$. The natural setting for these groups is within a Clifford algebra. And the action of the rotations is produced by a kind of "sandwich", denoted by qvq^*.

Infinitesimal rotations

The matrices in the Lie algebra are not themselves rotations; the skew-symmetric matrices are derivatives, proportional differences of rotations. An actual "differential rotation", or *infinitesimal rotation matrix* has the form

where $d\theta$ is vanishingly small. These matrices do not satisfy all the same properties as ordinary finite rotation matrices under the usual treatment of infinitesimals (Goldstein, Poole & Safko 2002, §4.8). To understand what

matrices under the usual treatment of infinitesimals (Goldstein, Poole & Safko 2002, §4.8). To understand what this means, consider

We first test the orthogonality condition, $Q^{T}Q = I$. The product is

differing from an identity matrix by second order infinitesimals, which we discard. So to first order, an infinitesimal rotation matrix is an orthogonal matrix. Next we examine the square of the matrix.

Again discarding second order effects, we see that the angle simply doubles. This hints at the most essential difference in behavior, which we can exhibit with the assistance of a second infinitesimal rotation,

Compare the products $dA_{\mathbf{x}}dA_{\mathbf{y}}$ and $dA_{\mathbf{y}}dA_{\mathbf{x}}$.

Since $d\theta\ d\phi$ is second order, we discard it; thus, to first order, multiplication of infinitesimal rotation matrices is commutative. In fact,

again to first order. Put in other words, **the order in which infinitesimal rotations are applied is irrelevant**, this useful fact makes, for example, derivation of rigid body rotation relatively simple.

But we must always be careful to distinguish (the first order treatment of) these infinitesimal rotation matrices from both finite rotation matrices and from derivatives of rotation matrices (namely skew-symmetric matrices). Contrast the behavior of finite rotation matrices in the BCH formula with that of infinitesimal rotation matrices, where all the commutator terms will be second order infinitesimals so we *do* have a vector space.

Conversions

Main article: Rotation representation (mathematics)#Conversion formulae between representations

We have seen the existence of several decompositions that apply in any dimension, namely independent planes, sequential angles, and nested dimensions. In all these cases we can either decompose a matrix or construct one. We have also given special attention to 3×3 rotation matrices, and these warrant further attention, in both directions (Stuelpnagel 1964).

Quaternion

Main article: Quaternions and spatial rotation

Given the unit quaternion $q = (w,x,y,z)$, the equivalent 3×3 rotation matrix is

Now every quaternion component appears multiplied by two in a term of degree two, and if all such terms are zero what's left is an identity matrix. This leads to an efficient, robust conversion from any quaternion — whether unit,

nonunit, or even zero — to a 3×3 rotation matrix.

```
Nq = w^2 + x^2 + y^2 + z^2
if Nq > 0.0 then s = 2/Nq else s = 0.0
X = x*s; Y = y*s; Z = z*s
wX = w*X; wY = w*Y; wZ = w*Z
xX = x*X; xY = x*Y; xZ = x*Z
yY = y*Y; yZ = y*Z; zZ = z*Z
[ 1.0-(yY+zZ)    xY-wZ        xZ+wY     ]
[     xY+wZ    1.0-(xX+zZ)    yZ-wX     ]
[     xZ-wY       yZ+wX    1.0-(xX+yY)  ]
```

Freed from the demand for a unit quaternion, we find that nonzero quaternions act as homogeneous coordinates for 3×3 rotation matrices. The Cayley transform, discussed earlier, is obtained by scaling the quaternion so that its w component is 1. For a 180° rotation around any axis, w will be zero, which explains the Cayley limitation.

The sum of the entries along the main diagonal (the trace), plus one, equals $4-4(x^2+y^2+z^2)$, which is $4w^2$. Thus we can write the trace itself as $2w^2+2w^2-1$; and from the previous version of the matrix we see that the diagonal entries themselves have the same form: $2x^2+2w^2-1$, $2y^2+2w^2-1$, and $2z^2+2w^2-1$. So we can easily compare the magnitudes of all four quaternion components using the matrix diagonal. We can, in fact, *obtain* all four magnitudes using sums and square roots, and choose consistent signs using the skew-symmetric part of the off-diagonal entries.

```
w = 0.5*sqrt(1+Qxx+Qyy+Qzz)
x = copysign(0.5*sqrt(1+Qxx-Qyy-Qzz),Qzy-Qyz)
y = copysign(0.5*sqrt(1-Qxx+Qyy-Qzz),Qxz-Qzx)
z = copysign(0.5*sqrt(1-Qxx-Qyy+Qzz),Qyx-Qxy)
```

where copysign(x,y) is x with the sign of y:

Alternatively, use a single square root and division

```
t = Qxx+Qyy+Qzz
r = sqrt(1+t)
s = 0.5/r
w = 0.5*r
x = (Qzy-Qyz)*s
y = (Qxz-Qzx)*s
z = (Qyx-Qxy)*s
```

This is numerically stable so long as the trace, `t`, is not negative; otherwise, we risk dividing by (nearly) zero. In that case, suppose Q_{xx} is the largest diagonal entry, so x will have the largest magnitude (the other cases are similar); then the following is safe.

```
r = sqrt(1+Qxx-Qyy-Qzz)
s = 0.5/r
w = (Qzy-Qyz)*s
x = 0.5*r
y = (Qxy+Qyx)*s
z = (Qzx+Qxz)*s
```

If the matrix contains significant error, such as accumulated numerical error, we may construct a symmetric 4×4 matrix,

and find the eigenvector, (w,x,y,z), of its largest magnitude eigenvalue. (If Q is truly a rotation matrix, that value will be 1.) The quaternion so obtained will correspond to the rotation matrix closest to the given matrix (Bar-Itzhack 2000).

Polar decomposition

If the $n \times n$ matrix M is non-singular, its columns are linearly independent vectors; thus the Gram–Schmidt process can adjust them to be an orthonormal basis. Stated in terms of numerical linear algebra, we convert M to an orthogonal matrix, Q, using QR decomposition. However, we often prefer a Q "closest" to M, which this method does not accomplish. For that, the tool we want is the polar decomposition (Fan & Hoffman 1955; Higham 1989).

To measure closeness, we may use any matrix norm invariant under orthogonal transformations. A convenient choice is the Frobenius norm, $\|Q-M\|_F$, squared, which is the sum of the squares of the element differences. Writing this in terms of the trace, Tr, our goal is,

- Find Q minimizing $\mathrm{Tr}(\ (Q-M)^T(Q-M)\ )$, subject to $Q^TQ = I$.

Though written in matrix terms, the objective function is just a quadratic polynomial. We can minimize it in the usual way, by finding where its derivative is zero. For a 3×3 matrix, the orthogonality constraint implies six scalar equalities that the entries of Q must satisfy. To incorporate the constraint(s), we may employ a standard technique, Lagrange multipliers, assembled as a symmetric matrix, Y. Thus our method is:

- Differentiate $\mathrm{Tr}(\ (Q-M)^T(Q-M) + (Q^TQ-I)Y\ )$ with respect to (the entries of) Q, and equate to zero.

In general, we obtain the equation

so that

where Q is orthogonal and S is symmetric. To ensure a minimum, the Y matrix (and hence S) must be positive definite. Linear algebra calls QS the polar decomposition of M, with S the positive square root of $S^2 = M^TM$.

> Consider a 2×2 example. Including constraints, we seek to minimize
>
> Taking the derivative with respect to Q_{xx}, Q_{xy}, Q_{yx}, Q_{yy} in turn, we assemble a matrix.

When M is non-singular, the Q and S factors of the polar decomposition are uniquely determined. However, the determinant of S is positive because S is positive definite, so Q inherits the sign of the determinant of M. That is, Q is only guaranteed to be orthogonal, not a rotation matrix. This is unavoidable; an M with negative determinant has no uniquely-defined closest rotation matrix.

Axis and angle

To efficiently construct a rotation matrix from an angle θ and a unit axis $\mathbf{u}$, we can take advantage of symmetry and skew-symmetry within the entries.

Axis and angle

To efficiently construct a rotation matrix from an angle θ and a unit axis **u**, we can take advantage of symmetry and skew-symmetry within the entries.

```
c = cos(θ); s = sin(θ); C = 1-c
xs = x*s;   ys = y*s;   zs = z*s
xC = x*C;   yC = y*C;   zC = z*C
xyC = x*yC; yzC = y*zC; zxC = z*xC
[ x*xC+c   xyC-zs   zxC+ys ]
[ xyC+zs   y*yC+c   yzC-xs ]
[ zxC-ys   yzC+xs   z*zC+c ]
```

Determining an axis and angle, like determining a quaternion, is only possible up to sign; that is, (**u**,θ) and (−**u**,−θ) correspond to the same rotation matrix, just like q and $-q$. As well, axis-angle extraction presents additional difficulties. The angle can be restricted to be from 0° to 180°, but angles are formally ambiguous by multiples of 360°. When the angle is zero, the axis is undefined. When the angle is 180°, the matrix becomes symmetric, which has implications in extracting the axis. Near multiples of 180°, care is needed to avoid numerical problems: in extracting the angle, a two-argument arctangent with `atan2(sin θ,cos θ)` equal to θ avoids the insensitivity of arccosine; and in computing the axis magnitude to force unit magnitude, a brute-force approach can lose accuracy through underflow (Moler & Morrison 1983).

A partial approach is as follows.

```
x = Qzy-Qyz
y = Qxz-Qzx
z = Qyx-Qxy
r = hypot(x,hypot(y,z))
t = Qxx+Qyy+Qzz
θ = atan2(r,t-1)
```

The x, y, and z components of the axis would then be divided by `r`. A fully robust approach will use different code when `t` is negative, as with quaternion extraction. When `r` is zero because the angle is zero, an axis must be provided from some source other than the matrix.

Euler angles

Complexity of conversion escalates with Euler angles (used here in the broad sense). The first difficulty is to establish which of the twenty-four variations of Cartesian axis order we will use. Suppose the three angles are θ_1, θ_2, θ_3; physics and chemistry may interpret these as

while aircraft dynamics may use

One systematic approach begins with choosing the right-most axis. Among all permutations of (x,y,z), only two place that axis first; one is an even permutation and the other odd. Choosing parity thus establishes the middle axis. That leaves two choices for the left-most axis, either duplicating the first or not. These three choices gives us 3×2×2 = 12 variations; we double that to 24 by choosing static or rotating axes.

One systematic approach begins with choosing the right-most axis. Among all permutations of (x,y,z), only two place that axis first; one is an even permutation and the other odd. Choosing parity thus establishes the middle axis. That leaves two choices for the left-most axis, either duplicating the first or not. These three choices gives us $3\times2\times2 = 12$ variations; we double that to 24 by choosing static or rotating axes.

This is enough to construct a matrix from angles, but triples differing in many ways can give the same rotation matrix. For example, suppose we use the **zyz** convention above; then we have the following equivalent pairs:

$$\begin{array}{rcll}
(90°, 45°, -105°) & \equiv & (-270°, -315°, 255°) & \textit{multiples of } 360° \\
(72°, 0°, 0°) & \equiv & (40°, 0°, 32°) & \textit{singular alignment} \\
(45°, 60°, -30°) & \equiv & (-135°, -60°, 150°) & \textit{bistable flip}
\end{array}$$

Angles for any order can be found using a concise common routine (Herter & Lott 1993; Shoemake 1994).

The problem of singular alignment, the mathematical analog of physical gimbal lock, occurs when the middle rotation aligns the axes of the first and last rotations. It afflicts every axis order at either even or odd multiples of 90°. These singularities are not characteristic of the rotation matrix as such, and only occur with the usage of Euler angles.

The singularities are avoided when considering and manipulating the rotation matrix as orthonormal row vectors (in 3D applications often named 'right'-vector, 'up'-vector and 'out'-vector) instead of as angles. The singularities are also avoided when working with quaternions.

Uniform random rotation matrices

We sometimes need to generate a uniformly distributed random rotation matrix. It seems intuitively clear in two dimensions that this means the rotation angle is uniformly distributed between 0 and 2π. That intuition is correct, but does not carry over to higher dimensions. For example, if we decompose 3×3 rotation matrices in axis-angle form, the angle should *not* be uniformly distributed; the probability that (the magnitude of) the angle is at most θ should be $^1/_\pi(\theta - \sin\theta)$, for $0 \le \theta \le \pi$.

Since SO(n) is a connected and locally compact Lie group, we have a simple standard criterion for uniformity, namely that the distribution be unchanged when composed with any arbitrary rotation (a Lie group "translation"). This definition corresponds to what is called *Haar measure*. León, Massé & Rivest (2006) show how to use the Cayley transform to generate and test matrices according to this criterion.

We can also generate a uniform distribution in any dimension using the *subgroup algorithm* of Diaconis & Shashahani (1987). This recursively exploits the nested dimensions group structure of SO(n), as follows. Generate a uniform angle and construct a 2×2 rotation matrix. To step from n to n+1, generate a vector **v** uniformly distributed on the n-sphere, S^n, embed the $n\times n$ matrix in the next larger size with last column (0,…,0,1), and rotate the larger matrix so the last column becomes **v**.

As usual, we have special alternatives for the 3×3 case. Each of these methods begins with three independent random scalars uniformly distributed on the unit interval. Arvo (1992) takes advantage of the odd dimension to change a Householder reflection to a rotation by negation, and uses that to aim the axis of a uniform planar rotation.

Another method uses unit quaternions. Multiplication of rotation matrices is homomorphic to multiplication of quaternions, and multiplication by a unit quaternion rotates the unit sphere. Since the homomorphism is a local isometry, we immediately conclude that to produce a uniform distribution on SO(3) we may use a uniform distribution on S^3.

Euler angles can also be used, though not with each angle uniformly distributed (Murnaghan 1962; Miles 1965).

For the axis-angle form, the axis is uniformly distributed over the unit sphere of directions, S^2, while the angle has the non-uniform distribution over [0,π] noted previously (Miles 1965).

See also

- Rotation representation
- Isometry
- Orthogonal matrix
- Rodrigues' rotation formula
- Yaw-pitch-roll system
- Plane of rotation

Notes

1. ^ W3C recommendation (2003). "Scalable Vector Graphics -- the initial coordinate system" (http://www.w3.org/TR/SVG/coords.html#InitialCoordinateSystem) . http://www.w3.org/TR/SVG/coords.html#InitialCoordinateSystem.
2. ^ Murray, Glenn (2005). "Rotation About an Arbitrary Axis in 3 Dimensions" (http://inside.mines.edu/~gmurray/ArbitraryAxisRotation/ArbitraryAxisRotation.html) . http://inside.mines.edu/~gmurray/ArbitraryAxisRotation/ArbitraryAxisRotation.html.

References

- Arvo, James (1992), "Fast random rotation matrices" (http://www.graphicsgems.org/) , in David Kirk, *Graphics Gems III*, San Diego: Academic Press Professional, pp. 117–120, ISBN 978-0-12-409671-4, http://www.graphicsgems.org/
- Baker, Andrew (2003), *Matrix Groups: An Introduction to Lie Group Theory*, Springer, ISBN 978-1-85233-470-3
- Bar-Itzhack, Itzhack Y. (Nov.–Dec. 2000), "New method for extracting the quaternion from a rotation matrix", *AIAA Journal of Guidance, Control and Dynamics* **23** (6): 1085–1087 (Engineering Note), doi:10.2514/2.4654 (http://dx.doi.org/10.2514%2F2.4654) , ISSN 0731-5090 (http://www.worldcat.org/issn/0731-5090)
- Björck, A.; Bowie, C. (June 1971), "An iterative algorithm for computing the best estimate of an orthogonal matrix", *SIAM Journal on Numerical Analysis* **8** (2): 358–364, doi:10.1137/0708036 (http://dx.doi.org/10.1137%2F0708036) , ISSN 0036-1429 (http://www.worldcat.org/issn/0036-1429)
- Cayley, Arthur (1846), "Sur quelques propriétés des déterminants gauches", *Journal für die Reine und Angewandte Mathematik (Crelle's Journal),* **32**: 119–123, ISSN 0075-4102 (http://www.worldcat.org/issn/0075-4102) ; reprinted as article 52 in Cayley, Arthur (1889), *The collected mathematical papers of Arthur Cayley* (http://www.hti.umich.edu/cgi/t/text/pageviewer-idx? c=umhistmath;cc=umhistmath;rgn=full%20text;idno=ABS3153.0001.001;didno=ABS3153.0001.001;view=i

c=umhistmath;cc=umhistmath;rgn=full%20text;idno=ABS3153.0001.001;didno=ABS3153.0001.001;view=i
mage;seq=00000349) , **I** (**1841–1853**), Cambridge University Press, pp. 332–336, http://www.hti.umich.edu/cgi/t/text/pageviewer-idx?c=umhistmath;cc=umhistmath;rgn=full%20text;idno=ABS3153.0001.001;didno=ABS3153.0001.001;view=image;seq=00000349

- Diaconis, Persi; Shahshahani, Mehrdad (1987), "The subgroup algorithm for generating uniform random variables", *Probability in the Engineering and Informational Sciences* **1**: 15–32, ISSN 0269-9648 (http://www.worldcat.org/issn/0269-9648)
- Engø, Kenth (June 2001), "On the BCH-formula in **so**(3)" (http://www.ii.uib.no/publikasjoner/texrap/abstract/2000-201.html) , *BIT Numerical Mathematics* **41** (3): 629–632, doi:10.1023/A:1021979515229 (http://dx.doi.org/10.1023%2FA%3A1021979515229) , ISSN 0006-3835 (http://www.worldcat.org/issn/0006-3835) , http://www.ii.uib.no/publikasjoner/texrap/abstract/2000-201.html
- Fan, Ky; Hoffman, Alan J. (February 1955), "Some metric inequalities in the space of matrices", *Proc. AMS* **6** (1): 111–116, doi:10.2307/2032662 (http://dx.doi.org/10.2307%2F2032662) , ISSN 0002-9939 (http://www.worldcat.org/issn/0002-9939)
- Fulton, William; Harris, Joe (1991), *Representation Theory: A First Course*, GTM, **129**, New York, Berlin, Heidelberg: Springer, MR1153249 (http://www.ams.org/mathscinet-getitem?mr=1153249) , ISBN 978-0-387-97495-8
- Goldstein, Herbert; Poole, Charles P.; Safko, John L. (2002), *Classical Mechanics* (third ed.), Addison Wesley, ISBN 978-0-201-65702-9
- Hall, Brian C. (2004), *Lie Groups, Lie Algebras, and Representations: An Elementary Introduction*, Springer, ISBN 978-0-387-40122-5 (GTM 222)
- Herter, Thomas; Lott, Klaus (September–October 1993), "Algorithms for decomposing 3-D orthogonal matrices into primitive rotations", *Computers & Graphics* **17** (5): 517–527, doi:10.1016/0097-8493(93)90003-R (http://dx.doi.org/10.1016%2F0097-8493%2893%2990003-R) , ISSN 0097-8493 (http://www.worldcat.org/issn/0097-8493)
- Higham, Nicholas J. (October 1 1989), "Matrix nearness problems and applications" (http://www.maths.manchester.ac.uk/~higham/pap-misc.html) , in Gover, M. J. C.; Barnett, S. (– Scholar search (http://scholar.google.co.uk/scholar?hl=en&lr=&q=author%3AHigham+intitle%3AApplications+of+Matrix+Theory&as_publication=&as_ylo=1989&as_yhi=1989&btnG=Search)), *Applications of Matrix Theory*, Oxford University Press, pp. 1–27, ISBN 978-0-19-853625-3, http://www.maths.manchester.ac.uk/~higham/pap-misc.html
- León, Carlos A.; Massé, Jean-Claude; Rivest, Louis-Paul (February 2006), "A statistical model for random rotations" (http://archimede.mat.ulaval.ca/pages/lpr/) , *Journal of Multivariate Analysis* **97** (2): 412–430, doi:10.1016/j.jmva.2005.03.009 (http://dx.doi.org/10.1016%2Fj.jmva.2005.03.009) , ISSN 0047-259X (http://www.worldcat.org/issn/0047-259X) , http://archimede.mat.ulaval.ca/pages/lpr/
- Miles, R. E. (December 1965), "On random rotations in R^3", *Biometrika* **52** (3/4): 636–639, doi:10.2307/2333716 (http://dx.doi.org/10.2307%2F2333716) , ISSN 0006-3444 (http://www.worldcat.org/issn/0006-3444)
- Moler, Cleve; Morrison, Donald (1983), "Replacing square roots by pythagorean sums" (http://domino.watson.ibm.com/tchjr/journalindex.nsf/0b9bc46ed06cbac1852565e6006fe1a0/0043d03ee1c1013c85256bfa0067f5a6?OpenDocument) , *IBM Journal of Research and Development* **27** (6), ISSN 0018-8646 (http://www.worldcat.org/issn/0018-8646) , http://domino.watson.ibm.com/tchjr/journalindex.nsf/0b9bc46ed06cbac1852565e6006fe1a0/0043d03ee1c1013c85256bfa0067f5a6?OpenDocument
- Murnaghan, Francis D. (1950), "The element of volume of the rotation group" (http://www.pnas.org/content/vol36/issue11/) , *Proceedings of the National Academy of Sciences* **36** (11): 670–672, doi:10.1073/pnas.36.11.670 (http://dx.doi.org/10.1073%2Fpnas.36.11.670) , ISSN 0027-8424 (http://www.worldcat.org/issn/0027-8424) , http://www.pnas.org/content/vol36/issue11/

e6923b0fa0007f3a0?OpenDocument

- Murnaghan, Francis D. (1950), "The element of volume of the rotation group" (http://www.pnas.org/content/vol36/issue11/) , *Proceedings of the National Academy of Sciences* **36** (11): 670–672, doi:10.1073/pnas.36.11.670 (http://dx.doi.org/10.1073%2Fpnas.36.11.670) , ISSN 0027-8424 (http://www.worldcat.org/issn/0027-8424) , http://www.pnas.org/content/vol36/issue11/
- Murnaghan, Francis D. (1962), *The Unitary and Rotation Groups*, Lectures on applied mathematics, Washington: Spartan Books
- Prentice, Michael J. (1986), "Orientation statistics without parametric assumptions", *Journal of the Royal Statistical Society. Series B (Methodological)* **48** (2): 214–222, ISSN 0035-9246 (http://www.worldcat.org/issn/0035-9246)
- Shepperd, Stanley W. (May–June 1978), "Quaternion from rotation matrix", *AIAA Journal of Guidance, Control and Dynamics* **1** (3): 223–224, ISSN 0731-5090 (http://www.worldcat.org/issn/0731-5090)
- Shoemake, Ken (1994), "Euler angle conversion" (http://www.graphicsgems.org/) , in Paul Heckbert, *Graphics Gems IV*, San Diego: Academic Press Professional, pp. 222–229, ISBN 978-0-12-336155-4, http://www.graphicsgems.org/
- Stuelpnagel, John (October 1964), "On the parameterization of the three-dimensional rotation group", *SIAM Review* **6** (4): 422–430, doi:10.1137/1006093 (http://dx.doi.org/10.1137%2F1006093) , ISSN 0036-1445 (http://www.worldcat.org/issn/0036-1445) (Also NASA-CR-53568 (http://ntrs.nasa.gov/search.jsp) .)
- Varadarajan, V. S. (1984), *Lie Groups, Lie Algebras, and Their Representation*, Springer, ISBN 978-0-387-90969-1 (GTM 102)
- Wedderburn, J. H. M. (1934) (– Scholar search (http://scholar.google.co.uk/scholar?hl=en&lr=&q=author%3AWedderburn+intitle%3ALectures+on+Matrices&as_publication=&as_ylo=1934&as_yhi=1934&btnG=Search)), *Lectures on Matrices* (http://www.ams.org/online_bks/coll17/) , AMS, ISBN 978-0-8218-3204-2, http://www.ams.org/online_bks/coll17/

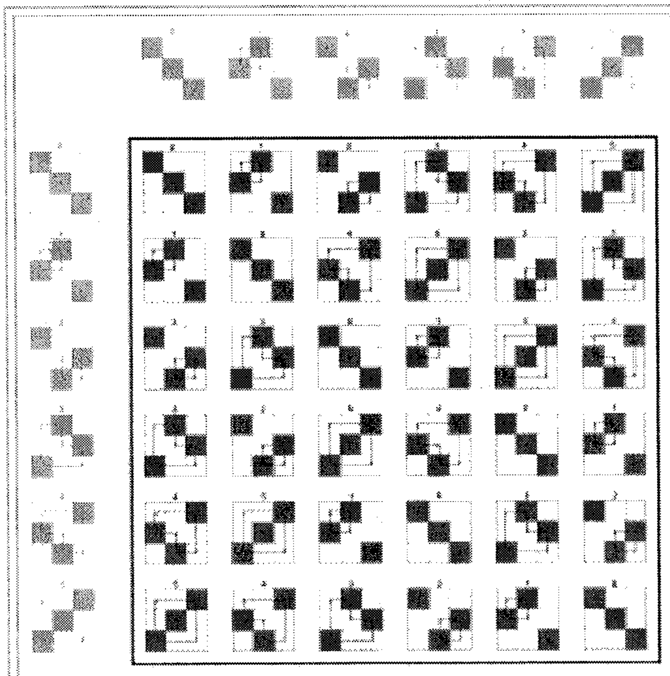
External links

- Rotation matrices at Mathworld (http://mathworld.wolfram.com/RotationMatrix.html)
- Math Awareness Month 2000 interactive demo (http://www.mathaware.org/mam/00/master/dimension/demos/plane-rotate.html) (requires Java)
- Rotation Matrices (http://www.mathpages.com/home/kmath593/kmath593.htm) at MathPages

Permutation matrix

From Wikipedia, the free encyclopedia

In mathematics, in matrix theory, a **permutation matrix** is a square binary matrix that has exactly one entry 1 in each row and each column and 0's elsewhere. Each such matrix represents a specific permutation of m elements and, when used to multiply another matrix, can produce that permutation in the rows or columns of the other matrix.



Matrices describing the permutations of 3 elements.
The product of two permutation matrices is a permutation matrix as well

Contents

Definition

Given a permutation π of m elements,

$$\pi : \{1,\ldots,m\} \rightarrow \{1,\ldots,m\}$$

given in two-line form by

$$\begin{pmatrix} 1 & 2 & \cdots & m \\ \pi(1) & \pi(2) & \cdots & \pi(m) \end{pmatrix},$$

its permutation matrix is the $m \times m$ matrix P_π whose entries are all 0 except that in row i, the entry $\pi(i)$ equals 1. We may write

$$P_\pi = \begin{bmatrix} \mathbf{e}_{\pi(1)} \\ \mathbf{e}_{\pi(2)} \\ \vdots \\ \mathbf{e}_{\pi(m)} \end{bmatrix},$$

where $\mathbf{e}_j$ denotes a row vector of length m with 1 in the jth position and 0 in every other position.

Properties

Given two permutations π and σ of m elements and the corresponding permutation matrices P_π and P_σ

$$P_\sigma P_\pi = P_{\pi \circ \sigma}$$

This somewhat unfortunate rule is a consequence of the definitions of multiplication of permutations (composition of bijections) and of matrices, and of the choice of using the vectors $\mathbf{e}_{\pi(i)}$ as rows of the permutation matrix; if one had used columns instead then the product above would have been equal to $P_{\sigma \circ \pi}$ with the permutations in their original order.

As permutation matrices are orthogonal matrices (i.e., $P_\pi P_\pi^T = I$), the inverse matrix exists and can be written as

$$P_\pi^{-1} = P_{\pi^{-1}} = P_\pi^T.$$

Multiplying P_π times a column vector $\mathbf{g}$ will permute the rows of the vector:

$$P_\pi \mathbf{g} = \begin{bmatrix} \mathbf{e}_{\pi(1)} \\ \mathbf{e}_{\pi(2)} \\ \vdots \\ \mathbf{e}_{\pi(n)} \end{bmatrix} \begin{bmatrix} g_1 \\ g_2 \\ \vdots \\ g_n \end{bmatrix} = \begin{bmatrix} g_{\pi(1)} \\ g_{\pi(2)} \\ \vdots \\ g_{\pi(n)} \end{bmatrix}.$$

Now applying P_σ after applying P_π gives the same result as applying $P_{\pi \circ \sigma}$ directly, in accordance with the above multiplication rule: call $P_\pi \mathbf{g} = \mathbf{g}'$, in other words

$$g'_i = g_{\pi(i)}$$

for all i, then

$$P_\sigma(P_\pi(\mathbf{g})) = P_\sigma(\mathbf{g}') = \begin{bmatrix} g'_{\sigma(1)} \\ g'_{\sigma(2)} \\ \vdots \\ g'_{\sigma(n)} \end{bmatrix} = \begin{bmatrix} g_{\pi(\sigma(1))} \\ g_{pi(\sigma(2))} \\ \vdots \\ g'_{\pi(\sigma(n))} \end{bmatrix}.$$

Multiplying a row vector **h** times P_π will permute the columns of the vector by the inverse of P_π:

$$\mathbf{h}P_\pi = \begin{bmatrix} h_1 \; h_2 \; \dots \; h_n \end{bmatrix} \begin{bmatrix} \mathbf{e}_{\pi(1)} \\ \mathbf{e}_{\pi(2)} \\ \vdots \\ \mathbf{e}_{\pi(n)} \end{bmatrix} = \begin{bmatrix} h_{\pi^{-1}(1)} \; h_{\pi^{-1}(2)} \; \dots \; h_{\pi^{-1}(n)} \end{bmatrix}$$

Again it can be checked that $(\mathbf{h}P_\sigma)P_\pi = \mathbf{h}P_{\pi\circ\sigma}$.

Notes

Let S_n denote the symmetric group, or group of permutations, on $\{1,2,...,n\}$. Since there are $n!$ permutations, there are $n!$ permutation matrices. By the formulas above, the $n \times n$ permutation matrices form a group under matrix multiplication with the identity matrix as the identity element.

If (1) denotes the identity permutation, then $P_{(1)}$ is the identity matrix.

One can view the permutation matrix of a permutation σ as the permutation σ of the columns of the identity matrix I, or as the permutation σ^{-1} of the rows of I.

A permutation matrix is a doubly stochastic matrix. The Birkhoff–von Neumann theorem says that every doubly stochastic matrix is a convex combination of permutation matrices of the same order and the permutation matrices are the extreme points of the set of doubly stochastic matrices. That is, the Birkhoff polytope, the set of doubly stochastic matrices, is the convex hull of the set of permutation matrices.

The product PM, premultiplying a matrix M by a permutation matrix P, permutes the rows of M; row i moves to row $\pi(i)$. Likewise, MP permutes the columns of M.

The map $S_n \to A \subset GL(n, \mathbf{Z}_2)$ is a faithful representation. Thus, $|A| = n!$.

The trace of a permutation matrix is the number of fixed points of the permutation. If the permutation has fixed points, so it can be written in cycle form as $\pi = (a_1)(a_2)...(a_k)\sigma$ where σ has no fixed points, then $\mathbf{e}_{a_1}, \mathbf{e}_{a_2}, ..., \mathbf{e}_{a_k}$ are eigenvectors of the permutation matrix.

From group theory we know that any permutation may be written as a product of transpositions. Therefore, any permutation matrix P factors as a product of row-interchanging elementary matrices, each having

determinant -1. Thus the determinant of a permutation matrix P is just the signature of the corresponding permutation.

Examples

The permutation matrix P_π corresponding to the permutation :$\pi = \begin{pmatrix} 1 & 2 & 3 & 4 & 5 \\ 1 & 4 & 2 & 5 & 3 \end{pmatrix}$, is

$$P_\pi = \begin{bmatrix} \mathbf{e}_{\pi(1)} \\ \mathbf{e}_{\pi(2)} \\ \mathbf{e}_{\pi(3)} \\ \mathbf{e}_{\pi(4)} \\ \mathbf{e}_{\pi(5)} \end{bmatrix} = \begin{bmatrix} \mathbf{e}_1 \\ \mathbf{e}_4 \\ \mathbf{e}_2 \\ \mathbf{e}_5 \\ \mathbf{e}_3 \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 \\ 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0 & 0 \end{bmatrix}.$$

Given a vector $\mathbf{g}$,

$$P_\pi \mathbf{g} = \begin{bmatrix} \mathbf{e}_{\pi(1)} \\ \mathbf{e}_{\pi(2)} \\ \mathbf{e}_{\pi(3)} \\ \mathbf{e}_{\pi(4)} \\ \mathbf{e}_{\pi(5)} \end{bmatrix} \begin{bmatrix} g_1 \\ g_2 \\ g_3 \\ g_4 \\ g_5 \end{bmatrix} = \begin{bmatrix} g_1 \\ g_4 \\ g_2 \\ g_5 \\ g_3 \end{bmatrix}.$$

Solving for *P*

If we are given two matrices A and B which are known to be related as $B = PAP^{-1}$, but the permutation matrix P itself is unknown, we can find P using eigenvalue decomposition:

$$A = Q_A \Lambda Q_A^{-1}$$
$$B = Q_B \Lambda Q_B^{-1}$$

where Λ is a diagonal matrix of eigenvalues, and Q_A and Q_B are the matrices of eigenvectors. The eigenvalues of A and B will always be the same, and P can be computed as $P = Q_B Q_A^T$. In other words, $PQ_A = Q_B$, which means that the eigenvectors of B are simply permuted eigenvectors of A.

[Note.The above processing of solving P is wrong. The right statement should be: there exist eigenvalue decomposition :$A = Q_A \Lambda Q_A^{-1}$

$B = Q_B \Lambda Q_B^{-1}$, such that $P = Q_B Q_A^T$ is a permutation solution of $B = PAP^{-1}$]

Example

Given the two matrices

$$A = \begin{bmatrix} 0 & 1 & 2 \\ 1 & 0 & 1.5 \\ 2 & 1.5 & 0 \end{bmatrix}$$
$$B = \begin{bmatrix} 0 & 1 & 1.5 \\ 1 & 0 & 2 \\ 1.5 & 2 & 0 \end{bmatrix}$$

and the transformation matrix that changes A into B is

$$P = \begin{bmatrix} 0 & 1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 1 \end{bmatrix}$$

which says that the first & second row as well as the first & second column of A have been swapped to yield B (and visual inspection confirms this).

After finding the eigenvalues of both A and B and diagonalizing them into a diagonal matrix is

$$\Lambda = \begin{bmatrix} -2.09394 & 0 & 0 \\ 0 & 0.9433954 & 0 \\ 0 & 0 & 3.037337 \end{bmatrix}$$

and the Q_A matrix of eigenvectors for A is

$$Q_A = \begin{bmatrix} -0.60130 & 0.54493 & 0.58437 \\ -0.25523 & -0.82404 & 0.50579 \\ 0.75716 & 0.15498 & 0.63458 \end{bmatrix}$$

and the Q_B matrix of eigenvectors for B is

$$Q_B = \begin{bmatrix} -0.25523 & -0.82404 & -0.50579 \\ -0.60130 & 0.54493 & -0.58437 \\ 0.75716 & 0.15498 & -0.63458 \end{bmatrix}.$$

Comparing the first eigenvector (i.e., the first column) of both we can write the first column of P by noting that the first element ($Q_{A(1,1)} = -0.60130$) matches the second element ($Q_{B(2,1)}$), thus we put a 1 in the second element of the first column of P. Repeating this procedure, we match the second element ($Q_{A(2,1)}$) to the first element ($Q_{B(1,1)}$), thus we put a 1 in the first element of the second column of P; and the third element ($Q_{A(3,1)}$) to the third element ($Q_{B(3,1)}$), thus we put a 1 in the third element of the third column of P.

The resulting P matrix is:

$$P = \begin{bmatrix} 0 & 1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 1 \end{bmatrix}.$$

And comparing to the P matrix from above, we find they are the same.

Explanation

A permutation matrix will always be in the form

$$\begin{bmatrix} \mathbf{e}_{a_1} \\ \mathbf{e}_{a_2} \\ \vdots \\ \mathbf{e}_{a_j} \end{bmatrix}$$

where $\mathbf{e}_{a_i}$ represents the ith basis vector (as a row) for $\mathbf{R}^j$, and where

$$\begin{bmatrix} 1 & 2 & \ldots & j \\ a_1 & a_2 & \ldots & a_j \end{bmatrix}$$

is the permutation form of the permutation matrix.

Now, in performing matrix multiplication, one essentially forms the dot product of each row of the first matrix with each column of the second. In this instance, we will be forming the dot product of each column of this matrix with the vector with elements we want to permute. That is, for example, = $(g_0,\ldots,g_5)^{\mathrm{T}}$,

$$\mathbf{e}_{a_i} \cdot \mathbf{v} = g_{a_i}$$

So, the product of the permutation matrix with the vector **v** above, will be a vector in the form $(g_{a_1}, g_{a_2}, \ldots, g_{a_j})$, and that this then is a permutation of **v** since we have said that the permutation form is

$$\begin{pmatrix} 1 & 2 & \ldots & j \\ a_1 & a_2 & \ldots & a_j \end{pmatrix}.$$

So, permutation matrices do indeed permute the order of elements in vectors multiplied with them.

Matrices with constant line sums

The sum of the values in each column or row in a permutation matrix adds up to exactly 1. A possible generalization of permutation matrices is nonnegative integral matrices where the values of each column and row add up to a constant number c. A matrix of this sort is known to be the sum of c permutation matrices.

For example in the following matrix M each column or row adds up to 5.

$$M = \begin{bmatrix} 5 & 0 & 0 & 0 & 0 \\ 0 & 3 & 2 & 0 & 0 \\ 0 & 0 & 0 & 5 & 0 \\ 0 & 1 & 2 & 0 & 2 \\ 0 & 1 & 1 & 0 & 3 \end{bmatrix}.$$

This matrix is the sum of 5 permutation matrices.

See also

- Alternating sign matrix
- Generalized permutation matrix

Retrieved from "http://en.wikipedia.org/wiki/Permutation_matrix"
Categories: Matrices | Permutations

- This page was last modified on 15 March 2010 at 12:14.